

Panel: Instructions – Language and Behavior

Bonnie Webber (Chairperson)

University of Pennsylvania
Philadelphia PA 19104-6389, USA
bonnie@central.cis.upenn.edu

Barbara Grosz

Harvard University
Cambridge MA 02138, USA
grosz@das.harvard.edu

Thomas Rist

German Research Center for AI
Saarbrücken, GERMANY
rist@dfki.uni-sb.de

Shigeoki Hirai

ETL
Tsukuba, JAPAN
hirai@etl.go.jp

Donia Scott

University of Brighton
Brighton BN2 4AT, UK
donia.scott@itri.bton.ac.uk

1 Panel Summary

This panel has three aims: (1) to inform the IJCAI-93 audience of the wide range of research in which instructions play a major role; (2) to explore some of the issues raised by trying to create effective instructions and/or to use them as intended; and (3) to link these issues with those of concern to mainstream AI. In short, we hope to make the audience aware of the benefits to be gained through work at the interface of *language* and *behavior*.

The panelists and their colleagues cover diverse parts of this range. (For pointers to other work, see Section 7.) **Grosz** and her colleagues are concerned with instructions in the context of multi-agent collaboration; **Hirai** and his colleagues are investigating instructions in support of telerobotic control; **Rist** and his colleagues at DFKI are concerned with the design and presentation of instructions using multiple media – creating and coordinating effective presentations automatically; **Scott** and her colleagues are addressing a major problem of the European Community – satisfying the need for quality instruction manuals in all the languages of the Community; **Webber** and colleagues at Penn are striving to use instructions to direct the behavior of animated human figures.

It is clear from the panelists' statements that there are several issues of general and immediate concern:

Representation how to represent the content of instructions, such that the representation is adequate for negotiating or guiding behavior, for generating multimedia presentations, or for producing quality texts in multiple languages;

Modelling how to model the agents receiving instructions, in order to choose instructions that they can understand and/or carry out as intended;

Performance Monitoring how to establish that the recipient of instructions has understood them and is following them appropriately and without error.

Instruction Quality how to ensure that instructions are coherent, intelligible and informative, and reflect what the recipient should indeed undertake to do.

This is the first panel we are aware of to address instructions from the viewpoint of AI, and we hope that it will benefit all those who take the time to listen or to read.

2 Instructions and Collaborative Plans

Barbara Grosz

Harvard University, Cambridge MA, USA

Cooperative problem solving by teams consisting of both humans and computers requires collaboration and communication. Our recent work has been directed at developing models of collaborative activity [GK93; GS90; LGS90] and at determining the types of information about actions and action relations provided by certain natural language constructions [Bal92b; Bal92a; Bal93]. Our model of collaborative plans, SharedPlans, specifies the beliefs and intentions needed for a group of agents to perform together a complex activity that decomposes into a combination of individual and multi-agent actions. Although originally intended to provide the basis for the intentional component of discourse structure [GS86], the model has been extended to cover multi-agent collaborative behavior more generally.

Instructions and SharedPlans have a two-way relationship. Knowledge of the way in which to perform an action is an important component of SharedPlans as well as of individual plans. We use the term *recipe* to refer to this type of knowledge; in particular, a recipe for α is a specification of a group of actions, which we will refer to as β_i , the doing of which under appropriate constraints, ρ_i , will constitute performance of α [Pol90; LGS90]. In this framework, instructions are a means by which one agent communicates recipe information to another. But also, this communication is an instance of collaborative behavior for which SharedPlans provide a

model.

The definition of SharedPlans [GK93] requires appeal to two kinds of intentions, *intentions to perform an action (Int.To)* and *intentions that a proposition holds (Int.Th)*. In a SharedPlan for α , the agent who will perform β_i has an Int.To perform β_i ; all other agents in the collaboration have an Int.Th this agent succeed. An "intention that" (Int.Th) leads to helpful behavior, but more specifically to an agent's intending to (Int.To) do an action γ if it believes its doing γ will help ensure the achievement of its intention that (Int.Th) the other agent succeed in doing β_i . The provision of instructions is one such γ .

Instructions may be provided under a range of interactive situations. The particulars of the situations affect the kinds of actions in which the instruction giver may engage to satisfy its Int.Th. For example, the instructions provided in a typical construction kit are a one-shot deal. The writer must anticipate all possible communicative needs, and address them in the instruction sheet. There is no opportunity for the writer's Int.Th to engender spontaneous new Int.To's. In contrast, instructions provided interactively while an agent is performing a task are an instance of full collaborative activity. The agents involved engage in a SharedPlan for communicating recipe information; new Int.Th's and Int.To's are generated continually.

We are currently examining the ways in which SharedPlans and individual plans interact in dialogue [Loc93], developing axioms of intending to and intending that, and investigating the interaction between SharedPlans and multi-media (natural language and graphics) communication [MR90].

References

- [Bal92a] C. T. Balkanski. Action relations in rationale clauses and means clauses. In *Proceedings of COLING-92*, Nantes, France, 1992.
- [Bal92b] C. T. Balkanski. Actions, beliefs and intentions in rationale clauses and means clauses. In *Proceedings of AAAI-92*, San Jose, CA, 1992.
- [Bal93] C. T. Balkanski. *Actions, Beliefs and Intentions in Multi-action Utterances*. PhD thesis, Harvard University, 1993. Forthcoming.
- [GK93] B. J. Grosz and S. Kraus. Collaborative plans for group activities. *Proceedings of IJCAI-93*, Chambéry, Savoie, France, 1993.
- [GS86] B.J. Grosz and C.L. Sidner. Attention, intentions, and the structure of discourse. *Computational Linguistics*, 12(3), 1986.
- [GS90] B.J. Grosz and C.L. Sidner. Plans for discourse. In P.R. Cohen, J.L. Morgan, and M.E. Pollack, editors, *Intentions in Communication*. MIT Press, 1990.
- [LGS90] K. E. Lochbaum, B. J. Grosz, and C. L. Sidner. Models of plans to support communication: An initial report. In *Proceedings of AAAI-90*, Boston, MA, 1990.
- [Loc93] K. E. Lochbaum. A collaborative planning approach to understanding subdialogues in conversation. Technical report, Harvard University, 1993. Forthcoming.
- [MR90] J. Marks and E. Reiter. Unwanted Conversational Implicatures in Text and Graphics In *Proceedings of the Eighth National Conference on Artificial Intelligence*, Boston, MA, 1990
- [Pol90] M. E. Pollack. Plans as complex mental attitudes. In P.R. Cohen, J.L. Morgan, and M.E. Pollack, editors, *Intentions in Communication*. MIT Press, 1990.

3 Language Instruction for Teleoperational Control

Shigeoki Hirai
ETL, Tsukuba, JAPAN

My colleague, Tomomasa Sato (Univ of Tokyo, JAPAN) and I have been conducting research on language instruction for teleoperational control of remote robot tasks. Key features of our approach are as follows: **Language-based master-slave manipulator system.** Our research started with the idea of introducing robot language to construct a master-slave manipulator system [SH87] to perform tasks such as chemical experiments in a hot cell. With this architecture, we hoped to combine the advantages of direct motion control with those of language specification. For example, the system could superpose language-specified motion constraints on direct control from the operation of the master-manipulator, thereby realizing human-robot cooperation at the motion level.

For example, "(SORI '(A B) '(0 0))" commands the robot to keep rotation angle of its hand around X and Y axis 0, which means that the robot keeps its hand horizontal. This function is useful when the robot carries a container filled with water. The operator has only to concentrate on where to put the container, while the robot takes care of keeping its posture unchanged so as not to spill the water.

Teleoperational task-oriented knowledge base. Even in teleoperational task execution, we need a variety of information about task objects, i.e. their handling properties and procedures for using them. It is quite natural to accumulate object handling knowledge that can be used when needed. It suits the aim of the telerobot, which is used for unexpected tasks in unstructured environments. We have constructed teleoperational task-oriented knowledge base [HS88].

For example, the system takes care of procedures related to a "pick-and-place" operation, including one for managing object-location data, calling teaching procedure if they are not filled yet, etc. This allows the operator to command the robot to operate an object, without the operator needing to check that all the necessary data is available beforehand or to update object-location data afterwards. The knowledge base also guides typical procedure for specific operations such as striking a match, igniting an lamp, etc. An appropriate teaching method

specialized to use shape features of the object is typical example of guidance from the knowledge base.

Language instruction as behavior models. A world model serves as a resource for task execution based on language instruction. But direct intervention through master-manipulator may bring about differences between contents of the manipulator's world model and the actual environment. We introduced motion understanding system to solve the problem [HS89]. In this system language instructions such as (**APPROACH object-x**) or (**GRIP object-x**) give conceptual frameworks of behavior model for telerobotic object handling. For example, the **GRIP** instruction can be used when the robot holds nothing in its hand and the object to be grasped is in the middle of its fingers. The robot can then close its fingers. In the definition of **GRIP** are rules that interpret the slave manipulator's motion guided by the operator, and the interpreted results are used to maintain world model automatically.

Instructing other teleoperational functions than manipulator. As a word has several meaning, so has an instruction to a telerobot. The issues that we addressed above are primarily concerned with telerobotic motion. Supervising a remote robot means not only giving it tasks but also observing its task progress and results. We have developed an intelligent camera work control for monitoring remote robot tasks along with the task execution [Wa92]. One interesting point here is that the same instruction has different effects on the robot and on the camera.

For example, instructing the robot (**APPROACH object-x**) makes the robot-hand approach the specified object. At the same time, it controls the camera to observe the entire path of robot motion, i.e. from current hand position to object's location. Thus the same instruction affects the devices in different ways.

Communication with telerobot through language. From the viewpoint of information flow between human and robot, research topics in telerobotics so far have concerned mainly with sending information to remote robots, through instructions from human to robot. We recognize that information flow from robot to human is also important. We expect robots to be able to report not only task specific information but also any other information about the remote environment, on the request of the human operator.

For example, if the robot falls into trouble in the midst of a task, the operator may want ask "What's the matter? What's the cause of the trouble?," and hope the robot to answer as "An unknown (not described in the program or world model) object obstructs my way. I cannot find how to remove it."

Communication with telerobot through language will be quite interesting research topic.

References

[SH87] Sato, T. and Hirai, S. Language-Aided Robotic Teleoperation System (LARTS) for Advanced Teleoperation, *IEEE J. Robotics and Automation* RA-3(5), pp.476-480, 1987.

[HS88] Hirai, S. and Sato, T. OBJECT MODEL FOR TELEROBOT, *Proc. IROS '88*, Tokyo, 1988, pp. 701-706.

[HS89] Hirai, S. and Sato, T. Motion Understanding for World Model Management of Telerobot, *Proc. ISRR*, Tokyo, 1989, pp.124-131.

[HS89] Hirai, S. et al. Intelligent and Cooperative Control of Telerobot Tasks, *IEEE Control Systems Magazine* 12(1), 1992, pp.51-56.

[Wa92] Wakita et al. Automatic Camera-work Control for Intelligent Monitoring of Telerobotic Tasks, *Proc. IROS92*, pp. 1130-1135, 1992.

4 Exploiting Multiple Media for Instructions

Thomas Rist Elisabeth Andre
DFKI, Saarbrücken GERMANY

There are several good reasons why most instruction manuals for technical devices appear as illustrated texts. Pictures show where parts are located and how things have to be manipulated. Pictures can also disambiguate referents via visually perceptible properties. Texts on the other hand enables us to abstract from irrelevant details and allows concepts to be conveyed, such as iteration and conditional action, that are hard to depict using static graphics alone.

Computer-based multimedia presentation techniques provide more effective means in instructing people in task performance as they are able to overcome problems arising from the non-adaptive, static and non-interactive nature of conventional technical documentation. However, instruction with multimedia presentation techniques is not advantageous per se. You might consider your own experience as regards badly crafted manuals in which illustrations may, rather than facilitating comprehension, even confuse readers.

Designing comprehensible instructions for human users that minimize execution time and errors is a challenging task. Questions that arise in the context of multimedia instructions include:

- What should be conveyed by text, what by graphics or other media?
- Which kinds of cohesive links between media are necessary?
- When should information be redundantly encoded by 'paraphrasing' it in different media?

A lot of work has already been done in diverse disciplines in order to give guidance to writers of technical manuals and instructions. To a certain extent, rules of thumb, such as 'use medium α for information β ', have been proven useful. However, in order to set up a sophisticated theory concerning how to design instructions, we have to spend much more attention to an instructee's *perceptual and cognitive processes*. For example, when selecting a presentation medium, one has not only to take into account the nature of the data and the medium,

but also the inferential processes a certain medium supports. As computer scientists we can offer new presentation forms as well as means for simulating both physical actions and cognitive processes. This includes both theoretical work on computational models and practical work on concrete systems.

The 'Intelligent User Interfaces' group at DFKI started in 1989 with the project WIP (Knowledge-based Presentation of Information) [AR92; Wa91; Wa93]. WIP aims to develop a presentation system that transforms formally represented domain plans in comprehensible instructions for human users by exploiting different presentation media such as text and graphics. The generation process is controlled by a set of generation parameters such as presentation objective, target group, resource limitations, and target language. The benefit of WIP lies in its ability to present the same information in a variety of ways depending on the value combination of the generation parameters.

The project PPP (Personalized Plan-Based Presenter) [An93] continues work done in the WIP project by adding three fundamental extensions. In contrast to WIP, the PPP system not only synthesizes multimedia documents, but also plans how this material is to be presented to various users. Furthermore, PPP allows for user interaction, i.e., PPP responds to follow-up questions concerning the domain as well as to meta comments on the act of presentation. Whereas WIP has no means of finding out whether a user really understood an instruction, PPP monitors the effectiveness of a presentation. The PPP system is able to keep track of the user's behavior and continuously adapts its instruction strategies to the current situation as it is based on a physical connection to the technical devices which are to be manipulated by the user.

References

- [An93] E. André, W. Graf, J. Heinsohn, B. Nebel, H.-J. Profitlich, T. Rist and W. Wahlster. *PPP- Personalized Plan-Based Presenter*. DFKI D-93-5, 1993.
- [AR92] E. André and T. Rist. The Design of Illustrated Documents as a Planning Task. In M. Maybury (ed.) *Intelligent Multimedia Interfaces*, AAAI-Press, 1993. (DFKI RR-92-45, 1993.)
- [Wa91] W. Wahlster, E. André, W. Graf and T. Rist. Designing Illustrated Texts. *Proc. of the 5th EACL*, pp. 8-14, Berlin, Germany, 1991.
- [Wa93] W. Wahlster, E. André, W. Finkler, H.-J. Profitlich and T. Rist. Plan-Based Integration of Natural Language and Graphics Generation. *AI Journal* 26(3). (DFKI RR-93-02, 1993.)

5 Multilingual Generation of Instructions

Donia R. Scott

University of Brighton, Brighton, UK

Instructions are an important application area within Europe, where there is an increasing demand for instructions for products and processes to be available in all the

languages of the Community. However, current practice for the production of commercial multilingual documentation hinders the efficient writing of effective instructions insofar as product/process development, authoring and translating tend to be three discrete and sequential processes, making use of different information and involving different skills:

domain experts Software product developers write product-oriented functional descriptions of the software system; developers of bureaucratic processes write texts written in a highly formulaic, legalistic language. Their skills are primarily technical/bureaucratic, not linguistic.

technical authors Professional writers of instructional documents use the information provided by the domain experts to write user-oriented instructions. Their skills are mainly linguistic and graphical rather than technical. They tend to be monolingual.

translators Translators work with the instructions produced by the technical writers and aim to produce an equally successful set of instructions appropriate to a different linguistic readership. Although translators tend to have similar skills to technical authors, they must be bilingual. While they should be good writers, they are not always so; they are measured primarily by their bi- or multilingual skills.

Because of the discrete nature of the processes, the final instructions tend to suffer from problems of drifting specifications¹ and of 'Chinese whispers'². Furthermore, the different specialisations of the two types of document producers very often lead to difficulties in producing good quality translations of the instructions: original texts are often (at best) badly written or (at worst) incomplete, misleading or even incomprehensible.

One approach to this problem is the use of a controlled language for the original instructions combined with the use of human or machine translation (the text-to-text approach).

We adopt an alternative approach—that of *multilingual generation* of instructions from a common representation of the underlying task. In two new collaborative projects involving academic colleagues in Austria, Italy, Spain and the UK, and groups of professional technical writers and translators of commercial and bureaucratic instructional texts, we are developing monolingual tools for (not necessarily bilingual) linguistic experts. The tools are to provide equivalent first drafts of instructional documents in selected languages of the European Community (English, French, Italian and German) and in two application areas (user manuals for software products, and instructions for bureaucratic procedures, e.g.

¹the lack of a common perspective among the three participant groups within the production process, combined with the lack of clear or complete initial information, resulting in documentation which does not meet the requirements of the user.

²the distortion and loss arising from misinterpretation, reduction or elision of the information that the translator receives from the technical author.

how to apply for a farm subsidy or for unemployment benefits).

A strong motivation for eschewing the text-to-text approach comes from the observation that texts are necessarily selective, embodying choices of expression appropriate only to the requirements of that one language; for the automatic generation of equivalent instructions, a linguistically predetermined representation cannot be the primary input, since the choices made for one language cannot, in general, be undone for the purposes of another.

Our multilingual drafting tools focus on two of the problems associated with the text-to-text approach for multilingual instructions [DHS92; DSH93]:

discourse structure Different languages often place different constraints on the presentation of information in instructional text. This is reflected in variations in clause ordering, discourse segment ordering and the ways that the relationships between clauses and segments are conveyed.

clause level considerations Different languages often impose markedly different requirements on the expression of any given fact. There are two sources of difficulty here. First, syntactic equivalence of constructions does not guarantee pragmatic equivalence. Second, different languages often do not require equivalent pragmatic markings in the expression of the same information.

The approach of multilingual generation provides the flexibility required to generate congruent, pragmatically equivalent, texts from a common underlying representations of the actions or procedures the reader is expected carry out. Clearly, this is important, since different languages often present instructions from different perspectives. A typical example of this can be seen in the following excerpts (taken from instructions for a step aerobics machine) which have the same underlying goal of getting the reader to balance the apparatus. Note that in the English and German versions the reader is informed *how* to balance the apparatus, while the French reader is told *why* the milled wheel needs to be turned. These different perspectives lead to different choices at the level of discourse structure, which in turn impact on choices at the clausal level.

The apparatus can be balanced on uneven floors by turning the milled wheel.

Durch Drehen der Rändelmutter kann das Gerät unebenen Bodenverhältnissen angeglichen werden.

By turning of the milled wheel can the apparatus (to) uneven floor conditions adapted become.

Tourner l'écrou molleté pour adapter l'appareil aux inégalités du plancher.
Turn the milled wheel to adapt the apparatus to the unevenness of the floor.

References

- [DHS92] J. Delin, A. Hartley and D.R. Scott. Comparative Pragmatics: Discourse Structure in Multilingual Document Generation. Internal Research Report, Information Technology Research Institute, University of Brighton, 1992.
- [DSH93] J. Delin, D.R. Scott and A. Hartley. Knowledge, Intention, Rhetoric: Levels of Variation in Multilingual Instructions. Submitted to the *ACL-93 Workshop on Intentionality and Structure in Discourse Relations*, Columbus OH, June 1993.

6 Instructions and Human Figure Animation

Bonnie Webber Norman Badler
U of Pennsylvania, Philadelphia PA, USA

The research being carried out in the *AnimNL* Project ("Animation and Natural Language") aims to enable people to use instructions to guide animated human figures through a task. The figures come from an animation system called *Jack*TM, developed by the Computer Graphics Research Lab at the University of Pennsylvania. *Jack* provides articulated, animated human figures capable of realistic motion through model-based behaviors [Ba93]. *Jack* agents can be anthropometrically sized and given different physical capabilities. In addition, different spatial environments can be constructed and modified at will, so as to vary the situations in which tasks are carried out.

Our ultimate goal is animated human figures that behave as semi-autonomous, instruction-understanding agents whose lowest-level behavioral goals can be fed into the simulator that sits above *Jack*, to be integrated and realized as realistic and informative animations. The potential uses of such a capability are many: from human task analysis, with tasks communicated through Natural Language, to small group training, where the other "members" of the group are animated human figures.

The problems are attempting to solve come at several different levels.

Language. The system must be able to interpret instructions in terms of (1) intentions that an agent should adopt, (2) desired constraints on the agent's behavior, and (3) expectations about what will happen. Doing this is not always straightforward. For example, Di Eugenio [Di92; DW92] has shown how purpose clauses affect what action an agent chooses to undertake: when told to "fold a square in half to create two triangles", an agent will perform a more specific action – folding the square along a diagonal – because only by doing so, will s/he create two triangles.

Planning. In response to instructions, the system must make symbolic decisions about how an agent should try to achieve the intentions it adopts. Again, this is not always straightforward. For example, when an agent is warned about something, it must adapt its behavior so as to comply. One way is by just avoiding the action being warned against. ("Don't leave your room.") But more complex types of warnings ("Be careful not to cut

outside the pattern lines", "Do not adjust the switch plunger to engage with maximum force", etc.) require more complex and indirect behavioral adjustments.

Behaviors in AnimNL must be achieved through an integration of incremental, hierarchical planning and reactive planning. The reactive planner is one that can directly "sense" the environment and control the locomotion of the agent directly. This obviates symbolic or global numerical path planning to approach a target, while avoiding collisions with nearby objects [BB93].

Two other planning problems that we have encountered are (1) that naturalistic behavior seems to demand that agents take account of future intentions when deciding how to satisfy current ones [Ge92; GW93]; (2) that objects are not always to hand or in known places when needed, so an agent must then look for them [Mo93].

Obviously, these are general issues of planning and rational behavior, independent of language, but the need to act appropriately in response to instructions makes the need for solutions more pressing.

Body. Some problems lie at the level of the agent's body – how an agent with a given body size and strength, and with the resources it currently has available, should move itself in order to achieve its intentions. For example, an agent must decide where to stand in order to open a door to let someone in, or how to position itself and bend in order to reach a ball that has rolled underneath a table [Ba93]. Even for apparently simple "grasp" actions, considerable reasoning is needed: gaining control over an accessible object requires that one determine an appropriate approach path, grasp, and grasp site to achieve task intentions.

Integration. Finally, some problems have to do with how levels are connected such that current and future intentions can propagate *downwards* and be used to generate natural body movements, while performance inability or failure can propagate *upwards* to affect subsequent planning and re-planning decisions. Coordinating the performance of multiple simultaneous and sequential actions through simulation is a non-trivial control problem. As for performance failure, whether they can be handled at the motor control, reactive planning, or instruction understanding level may depend on other factors. In particular, we must understand the influence of the agent's inherent skills, the coordination of multiple simultaneous task actions, the role of previous task performance, the process of tool selection and use, and the use of assistance from other agents.

Contributors to the AnimNL Project include Tripp Becket, Barbara Di Eugenio, Chris Geib, Moon Jung, Libby Levison, Michael Moore, Michael White and Xinmin Zhao.

References

- [BB93] Becket, W. and Badler, N. Integrated Behavioral Agent Architecture. *Proc. 3rd Conference on Computer Generated Forces and Behavior Representation*, Orlando, FL, March 1993, pp.57-68.
 - [Ba93] Badler, N., Phillips, C. and Webber, B. *Simulated Agents and Virtual Humans*. Oxford University Press, 1993.
 - [Di92] Di Eugenio, B. Understanding Natural Language Instructions: the Case of Purpose Clauses. *Proc. 30th Annual Conference of the Assoc. for Computational Linguistics*, Newark DL, June 1992.
 - [DW92] Di Eugenio, B. and Webber, B. Plan Recognition in Understanding Instructions. *Proc. 1st. Int'l Conference on Artificial Intelligence Planning Systems*, College Park MD, June 1992, pp. 52–61.
 - [Ge92] Geib, C. *Intentions in Means/End Planning*. Technical Report MS-CIS-92-73, Dept. Computer & Information Science, Univ of Pennsylvania, 1992.
 - [GW93] Geib, C. and Webber, B. Alternatives to Preconditions in Means-End Planning. *AAAI Spring 1993 Symposium on Foundations of Automatic Planning*, Stanford CA, March 1993.
 - [Mo93] Moore, M.B. *Search Plans*. PhD proposal, Dept of Computer & Information Science, Univ of Pennsylvania, March 1993.
 - [We92] Webber, B., Badler, N., Baldwin, F.B., Becket, W., DiEugenio, B., Geib, C., Jung, M., Levison, L. and Moore, M. Doing What You're Told. Tech Report MS-CIS-92-74, Dept of Computer & Information Science, Univ of Pennsylvania, Sept 1992.
- ## 7 Additional References
- [AC90] Agre, P. and Chapman, D. What are Plans for? In P. Maes (ed.), *Designing Autonomous Agents*, Cambridge MA: MIT Press, 1990.
 - [Al91] Alterman, R., Zito-Wolf, R. and Carpenter, T. Interaction, Comprehension, and Instruction Usage. *J. Learning Sciences* 1(4), 1991.
 - [Ch91] Chapman, D. *Vision, Instruction and Action*. Cambridge: MIT Press, 1991.
 - [Cr89] Crangle, C. On Saying 'Stop' to a Robot. *Language and Communication* 9(1), 1989.
 - [Da92] Dale, R. *Generating Referring Expressions: Constructing Descriptions in a Domain of Objects and Processes*, Cambridge MA: MIT Press, 1992.
 - [FM91] Feiner, S. and McKeown, K. Automating the Generation of Coordinated Multimedia Explanations. *IEEE Computer*, 24(10), October 1991.
 - [HL93] Huffman, S. and Laird, J. Learning Procedures from Interactive Natural Language Instructions. *Proc. 10th Int'l Conf. on Machine Learning*, Amherst MA, June 1993.
 - [Mc92] McKeown, K., Feiner, S., Robin, J., Seligmann, D., and Tanenblatt, M. Generating Cross-References for Multimedia Explanation. *Proc. AAAI-92*, San Jose CA, July 1992.
 - [VB90] Vere, S. and Bickmore, T. A basic agent. *Computational Intelligence*, 6(1), 1990.